

RECEIVED

APR 01 2003

TECH CENTER 1600/2500



1600

RAW SEQUENCE LISTING

DATE: 03/25/2003

PATENT APPLICATION: US/09/254,344A

TIME: 12:53:43

Input Set : A:\024705-077.ST25.txt

Output Set: N:\CRF4\03252003\I254344A.raw

4 <110> APPLICANT: HAYASHIZAKI, Yoshihide
 5 WATAHIKI, Masanori
 7 <120> TITLE OF INVENTION: RNA Polymerase
 9 <130> FILE REFERENCE: 024705-077
 11 <140> CURRENT APPLICATION NUMBER: US 09/254,344A
 12 <141> CURRENT FILING DATE: 1999-09-03
 14 <150> PRIOR APPLICATION NUMBER: PCT/JP98/03037
 15 <151> PRIOR FILING DATE: 1998-07-06
 17 <150> PRIOR APPLICATION NUMBER: JP 9/180883
 18 <151> PRIOR FILING DATE: 1997-07-07
 20 <150> PRIOR APPLICATION NUMBER: JP 10/155759
 21 <151> PRIOR FILING DATE: 1998-06-04
 23 <160> NUMBER OF SEQ ID NOS: 39
 25 <170> SOFTWARE: PatentIn version 3.0
 27 <210> SEQ ID NO: 1
 28 <211> LENGTH: 2659
 29 <212> TYPE: DNA
 30 <213> ORGANISM: Bacteriophage T7
 32 <220> FEATURE:
 33 <221> NAME/KEY: CDS
 34 <222> LOCATION: (10)..(2658)
 36 <400> SEQUENCE: 1

37	aggcactaa atg aac acg att aac atc gct aag aac gac ttc tct gac atc	51
38	Met Asn Thr Ile Asn Ile Ala Lys Asn Asp Phe Ser Asp Ile	
39	1 5 10	
41	gaa ctg gct gct atc ccg ttc aac act ctg gct gac cat tac ggt gag	99
42	Glu Leu Ala Ala Ile Pro Phe Asn Thr Leu Ala Asp His Tyr Gly Glu	
43	15 20 25 30	
45	cgt tta gct cgc gaa cag ttg gcc ctt gag cat gag tct tac gag atg	147
46	Arg Leu Ala Arg Glu Gln Leu Ala Leu Glu His Glu Ser Tyr Glu Met	
47	35 40 45	
49	ggt gaa gca cgc ttc cgc aag atg ttt gag cgt caa ctt aaa gct ggt	195
50	Gly Glu Ala Arg Phe Arg Lys Met Phe Glu Arg Gln Leu Lys Ala Gly	
51	50 55 60	
53	gag gtt gcg gat aac gct gcc gcc aag cct ctc atc act acc cta ctc	243
54	Glu Val Ala Asp Asn Ala Ala Ala Lys Pro Leu Ile Thr Thr Leu Leu	
55	65 70 75	
57	cct aag atg att gca cgc atc aac gac tgg ttt gag gaa gtg aaa gct	291
58	Pro Lys Met Ile Ala Arg Ile Asn Asp Trp Phe Glu Glu Val Lys Ala	
59	80 85 90	
61	aag cgc ggc aag cgc ccg aca gcc ttc cag ttc ctg caa gaa atc aag	339
62	Lys Arg Gly Lys Arg Pro Thr Ala Phe Gln Phe Leu Gln Glu Ile Lys	
63	95 100 105 110	

P.6
ENTERED

RAW SEQUENCE LISTING

DATE: 03/25/2003

PATENT APPLICATION: US/09/254,344A

TIME: 12:53:43

Input Set : A:\024705-077.ST25.txt

Output Set: N:\CRF4\03252003\I254344A.raw

```

65 ccg gaa gcc gta gcg tac atc acc att aag acc act ctg gct tgc cta      387
66 Pro Glu Ala Val Ala Tyr Ile Thr Ile Lys Thr Thr Leu Ala Cys Leu
67          115          120          125
69 acc agt gct gac aat aca acc gtt cag gct gta gca agc gca atc ggt      435
70 Thr Ser Ala Asp Asn Thr Thr Val Gln Ala Val Ala Ser Ala Ile Gly
71          130          135          140
73 cgg gcc att gag gac gag gct cgc ttc ggt cgt atc cgt gac ctt gaa      483
74 Arg Ala Ile Glu Asp Glu Ala Arg Phe Gly Arg Ile Arg Asp Leu Glu
75          145          150          155
77 gct aag cac ttc aag aaa aac gtt gag gaa caa ctc aac aag cgc gta      531
78 Ala Lys His Phe Lys Lys Asn Val Glu Glu Gln Leu Asn Lys Arg Val
79          160          165          170
81 ggg cac gtc tac aag aaa gca ttt atg caa gtt gtc gag gct gac atg      579
82 Gly His Val Tyr Lys Lys Ala Phe Met Gln Val Val Glu Ala Asp Met
83 175          180          185          190
85 ctc tct aag ggt cta ctc ggt ggc gag gcg tgg tct tcg tgg cat aag      627
86 Leu Ser Lys Gly Leu Leu Gly Gly Glu Ala Trp Ser Ser Trp His Lys
87          195          200          205
89 gaa gac tct att cat gta gga gta cgc tgc atc gag atg ctc att gag      675
90 Glu Asp Ser Ile His Val Gly Val Arg Cys Ile Glu Met Leu Ile Glu
91          210          215          220
93 tca acc gga atg gtt agc tta cac cgc caa aat gct ggc gta gta ggt      723
94 Ser Thr Gly Met Val Ser Leu His Arg Gln Asn Ala Gly Val Val Gly
95          225          230          235
97 caa gac tct gag act atc gaa ctc gca cct gaa tac gct gag gct atc      771
98 Gln Asp Ser Glu Thr Ile Glu Leu Ala Pro Glu Tyr Ala Glu Ala Ile
99          240          245          250
101 gca acc cgt gca ggt gcg ctg gct ggc atc tct ccg atg ttc caa cct      819
102 Ala Thr Arg Ala Gly Ala Leu Ala Gly Ile Ser Pro Met Phe Gln Pro
103 255          260          265          270
105 tgc gta gtt cct cct aag ccg tgg act ggc att act ggt ggt ggc tat      867
106 Cys Val Val Pro Pro Lys Pro Trp Thr Gly Ile Thr Gly Gly Gly Tyr
107          275          280          285
109 tgg gct aac ggt cgt cgt cct ctg gcg ctg gtg cgt act cac agt aag      915
110 Trp Ala Asn Gly Arg Arg Pro Leu Ala Leu Val Arg Thr His Ser Lys
111          290          295          300
113 aaa gca ctg atg cgc tac gaa gac gtt tac atg cct gag gtg tac aaa      963
114 Lys Ala Leu Met Arg Tyr Glu Asp Val Tyr Met Pro Glu Val Tyr Lys
115          305          310          315
117 gcg att aac att gcg caa aac acc gca tgg aaa atc aac aag aaa gtc      1011
118 Ala Ile Asn Ile Ala Gln Asn Thr Ala Trp Lys Ile Asn Lys Lys Val
119          320          325          330
121 cta gcg gtc gcc aac gta atc acc aag tgg aag cat tgt ccg gtc gag      1059
122 Leu Ala Val Ala Asn Val Ile Thr Lys Trp Lys His Cys Pro Val Glu
123 335          340          345          350
125 gac atc cct gcg att gag cgt gaa gaa ctc ccg atg aaa ccg gaa gac      1107
126 Asp Ile Pro Ala Ile Glu Arg Glu Glu Leu Pro Met Lys Pro Glu Asp
127          355          360          365
129 atc gac atg aat cct gag gct ctc acc gcg tgg aaa cgt gct gcc gct      1155

```

RAW SEQUENCE LISTING

DATE: 03/25/2003

PATENT APPLICATION: US/09/254,344A

TIME: 12:53:43

Input Set : A:\024705-077.ST25.txt

Output Set: N:\CRF4\03252003\I254344A.raw

```

130 Ile Asp Met Asn Pro Glu Ala Leu Thr Ala Trp Lys Arg Ala Ala Ala
131           370           375           380
133 gct gtg tac cgc aag gac aag gct cgc aag tct cgc cgt atc agc ctt      1203
134 Ala Val Tyr Arg Lys Asp Lys Ala Arg Lys Ser Arg Arg Ile Ser Leu
135           385           390           395
137 gag ttc atg ctt gag caa gcc aat aag ttt gct aac cat aag gcc atc      1251
138 Glu Phe Met Leu Glu Gln Ala Asn Lys Phe Ala Asn His Lys Ala Ile
139           400           405           410
141 tgg ttc cct tac aac atg gac tgg cgc ggt cgt gtt tac gct gtg tca      1299
142 Trp Phe Pro Tyr Asn Met Asp Trp Arg Gly Arg Val Tyr Ala Val Ser
143 415           420           425           430
145 atg ttc aac ccg caa ggt aac gat atg acc aaa gga ctg ctt acg ctg      1347
146 Met Phe Asn Pro Gln Gly Asn Asp Met Thr Lys Gly Leu Leu Thr Leu
147           435           440           445
149 gcg aaa ggt aaa cca atc ggt aag gaa ggt tac tac tgg ctg aaa atc      1395
150 Ala Lys Gly Lys Pro Ile Gly Lys Glu Gly Tyr Tyr Trp Leu Lys Ile
151           450           455           460
153 cac ggt gca aac tgt gcg ggt gtc gat aag gtt ccg ttc cct gag cgc      1443
154 His Gly Ala Asn Cys Ala Gly Val Asp Lys Val Pro Phe Pro Glu Arg
155           465           470           475
157 atc aag ttc att gag gaa aac cac gag aac atc atg gct tgc gct aag      1491
158 Ile Lys Phe Ile Glu Glu Asn His Glu Asn Ile Met Ala Cys Ala Lys
159           480           485           490
161 tct cca ctg gag aac act tgg tgg gct gag caa gat tct ccg ttc tgc      1539
162 Ser Pro Leu Glu Asn Thr Trp Trp Ala Glu Gln Asp Ser Pro Phe Cys
163 495           500           505           510
165 ttc ctt gcg ttc tgc ttt gag tac gct ggg gta cag cac cac ggc ctg      1587
166 Phe Leu Ala Phe Cys Phe Glu Tyr Ala Gly Val Gln His His Gly Leu
167           515           520           525
169 agc tat aac tgc tcc ctt ccg ctg gcg ttt gac ggg tct tgc tct ggc      1635
170 Ser Tyr Asn Cys Ser Leu Pro Leu Ala Phe Asp Gly Ser Cys Ser Gly
171           530           535           540
173 atc cag cac ttc tcc gcg atg ctc cga gat gag gta ggt ggt cgc gcg      1683
174 Ile Gln His Phe Ser Ala Met Leu Arg Asp Glu Val Gly Gly Arg Ala
175           545           550           555
177 gtt aac ttg ctt cct agt gaa acc gtt cag gac atc tac ggg att gtt      1731
178 Val Asn Leu Leu Pro Ser Glu Thr Val Gln Asp Ile Tyr Gly Ile Val
179           560           565           570
181 gct aag aaa gtc aac gag att cta caa gca gac gca atc aat ggg acc      1779
182 Ala Lys Lys Val Asn Glu Ile Leu Gln Ala Asp Ala Ile Asn Gly Thr
183 575           580           585           590
185 gat aac gaa gta gtt acc gtg acc gat gag aac act ggt gaa atc tct      1827
186 Asp Asn Glu Val Val Thr Val Thr Asp Glu Asn Thr Gly Glu Ile Ser
187           595           600           605
189 gag aaa gtc aag ctg ggc act aag gca ctg gct ggt caa tgg ctg gct      1875
190 Glu Lys Val Lys Leu Gly Thr Lys Ala Leu Ala Gly Gln Trp Leu Ala
191           610           615           620
193 tac ggt gtt act cgc agt gtg act aag cgt tca gtc atg acg ctg gct      1923
194 Tyr Gly Val Thr Arg Ser Val Thr Lys Arg Ser Val Met Thr Leu Ala

```

RAW SEQUENCE LISTING

DATE: 03/25/2003

PATENT APPLICATION: US/09/254,344A

TIME: 12:53:43

Input Set : A:\024705-077.ST25.txt

Output Set: N:\CRF4\03252003\I254344A.raw

195	625	630	635	
197	tac ggg tcc aaa gag ttc ggc ttc cgt caa caa gtg ctg gaa gat acc	1971		
198	Tyr Gly Ser Lys Glu Phe Gly Phe Arg Gln Gln Val Leu Glu Asp Thr			
199	640 645 650			
201	att cag cca gct att gat tcc ggc aag ggt ctg atg ttc act cag ccg	2019		
202	Ile Gln Pro Ala Ile Asp Ser Gly Lys Gly Leu Met Phe Thr Gln Pro			
203	655 660 665 670			
205	aat cag gct gct gga tac atg gct aag ctg att tgg gaa tct gtg agc	2067		
206	Asn Gln Ala Ala Gly Tyr Met Ala Lys Leu Ile Trp Glu Ser Val Ser			
207	675 680 685			
209	gtg acg gtg gta gct gcg gtt gaa gca atg aac tgg ctt aag tct gct	2115		
210	Val Thr Val Val Ala Ala Val Glu Ala Met Asn Trp Leu Lys Ser Ala			
211	690 695 700			
213	gct aag ctg ctg gct gct gag gtc aaa gat aag aag act gga gag att	2163		
214	Ala Lys Leu Leu Ala Ala Glu Val Lys Asp Lys Lys Thr Gly Glu Ile			
215	705 710 715			
217	ctt cgc aag cgt tgc gct gtg cat tgg gta act cct gat ggt ttc cct	2211		
218	Leu Arg Lys Arg Cys Ala Val His Trp Val Thr Pro Asp Gly Phe Pro			
219	720 725 730			
221	gtg tgg cag gaa tac aag aag cct att cag acg cgc ttg aac ctg atg	2259		
222	Val Trp Gln Glu Tyr Lys Lys Pro Ile Gln Thr Arg Leu Asn Leu Met			
223	735 740 745 750			
225	ttc ctc ggt cag ttc cgc tta cag cct acc att aac acc aac aaa gat	2307		
226	Phe Leu Gly Gln Phe Arg Leu Gln Pro Thr Ile Asn Thr Asn Lys Asp			
227	755 760 765			
229	agc gag att gat gca cac aaa cag gag tct ggt atc gct cct aac ttt	2355		
230	Ser Glu Ile Asp Ala His Lys Gln Glu Ser Gly Ile Ala Pro Asn Phe			
231	770 775 780			
233	gta cac agc caa gac ggt agc cac ctt cgt aag act gta gtg tgg gca	2403		
234	Val His Ser Gln Asp Gly Ser His Leu Arg Lys Thr Val Val Trp Ala			
235	785 790 795			
237	cac gag aag tac gga atc gaa tct ttt gca ctg att cac gac tcc ttc	2451		
238	His Glu Lys Tyr Gly Ile Glu Ser Phe Ala Leu Ile His Asp Ser Phe			
239	800 805 810			
241	ggt acc att ccg gct gac gct gcg aac ctg ttc aaa gca gtg cgc gaa	2499		
242	Gly Thr Ile Pro Ala Asp Ala Ala Asn Leu Phe Lys Ala Val Arg Glu			
243	815 820 825 830			
245	act atg gtt gac aca tat gag tct tgt gat gta ctg gct gat ttc tac	2547		
246	Thr Met Val Asp Thr Tyr Glu Ser Cys Asp Val Leu Ala Asp Phe Tyr			
247	835 840 845			
249	gac cag ttc gct gac cag ttg cac gag tct caa ttg gac aaa atg cca	2595		
250	Asp Gln Phe Ala Asp Gln Leu His Glu Ser Gln Leu Asp Lys Met Pro			
251	850 855 860			
253	gca ctt ccg gct aaa ggt aac ttg aac ctc cgt gac atc tta gag tcg	2643		
254	Ala Leu Pro Ala Lys Gly Asn Leu Asn Leu Arg Asp Ile Leu Glu Ser			
255	865 870 875			
257	gac ttc gcg ttc gcg t	2659		
258	Asp Phe Ala Phe Ala			
259	880			

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/254,344A

DATE: 03/25/2003

TIME: 12:53:43

Input Set : A:\024705-077.ST25.txt

Output Set: N:\CRF4\03252003\I254344A.raw

```

262 <210> SEQ ID NO: 2
263 <211> LENGTH: 883
264 <212> TYPE: PRT
265 <213> ORGANISM: Bacteriophage T7
267 <400> SEQUENCE: 2
268 Met Asn Thr Ile Asn Ile Ala Lys Asn Asp Phe Ser Asp Ile Glu Leu
269 1 5 10 15
272 Ala Ala Ile Pro Phe Asn Thr Leu Ala Asp His Tyr Gly Glu Arg Leu
273 20 25 30
276 Ala Arg Glu Gln Leu Ala Leu Glu His Glu Ser Tyr Glu Met Gly Glu
277 35 40 45
280 Ala Arg Phe Arg Lys Met Phe Glu Arg Gln Leu Lys Ala Gly Glu Val
281 50 55 60
284 Ala Asp Asn Ala Ala Ala Lys Pro Leu Ile Thr Thr Leu Leu Pro Lys
285 65 70 75 80
288 Met Ile Ala Arg Ile Asn Asp Trp Phe Glu Glu Val Lys Ala Lys Arg
289 85 90 95
292 Gly Lys Arg Pro Thr Ala Phe Gln Phe Leu Gln Glu Ile Lys Pro Glu
293 100 105 110
296 Ala Val Ala Tyr Ile Thr Ile Lys Thr Thr Leu Ala Cys Leu Thr Ser
297 115 120 125
300 Ala Asp Asn Thr Thr Val Gln Ala Val Ala Ser Ala Ile Gly Arg Ala
301 130 135 140
304 Ile Glu Asp Glu Ala Arg Phe Gly Arg Ile Arg Asp Leu Glu Ala Lys
305 145 150 155 160
308 His Phe Lys Lys Asn Val Glu Glu Gln Leu Asn Lys Arg Val Gly His
309 165 170 175
312 Val Tyr Lys Lys Ala Phe Met Gln Val Val Glu Ala Asp Met Leu Ser
313 180 185 190
316 Lys Gly Leu Leu Gly Gly Glu Ala Trp Ser Ser Trp His Lys Glu Asp
317 195 200 205
320 Ser Ile His Val Gly Val Arg Cys Ile Glu Met Leu Ile Glu Ser Thr
321 210 215 220
324 Gly Met Val Ser Leu His Arg Gln Asn Ala Gly Val Val Gly Gln Asp
325 225 230 235 240
328 Ser Glu Thr Ile Glu Leu Ala Pro Glu Tyr Ala Glu Ala Ile Ala Thr
329 245 250 255
332 Arg Ala Gly Ala Leu Ala Gly Ile Ser Pro Met Phe Gln Pro Cys Val
333 260 265 270
336 Val Pro Pro Lys Pro Trp Thr Gly Ile Thr Gly Gly Gly Tyr Trp Ala
337 275 280 285
340 Asn Gly Arg Arg Pro Leu Ala Leu Val Arg Thr His Ser Lys Lys Ala
341 290 295 300
344 Leu Met Arg Tyr Glu Asp Val Tyr Met Pro Glu Val Tyr Lys Ala Ile
345 305 310 315 320
348 Asn Ile Ala Gln Asn Thr Ala Trp Lys Ile Asn Lys Lys Val Leu Ala
349 325 330 335
352 Val Ala Asn Val Ile Thr Lys Trp Lys His Cys Pro Val Glu Asp Ile
353 340 345 350

```

RAW SEQUENCE LISTING ERROR SUMMARY
PATENT APPLICATION: US/09/254,344A

DATE: 03/25/2003
TIME: 12:53:44

Input Set : A:\024705-077.ST25.txt
Output Set: N:\CRF4\03252003\I254344A.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:17; N Pos. 25

Seq#:18; N Pos. 6,7,8,18,25,30

Seq#:20; N Pos. 6

Seq#:23; N Pos. 1,2,3,4,5,6,7,8,9,10,11,12,13,14,17,18,21,26,29,34,38,40,50

Seq#:23; N Pos. 51,53,58,75,76,85,110,117,132,142,150,157,399,440,472